

## **RAW SEQUENCE LISTING**

**Loaded by SCORE, no errors detected.**

Application Serial Number: 10/820,843

Source: 01PE

Date Processed by SCORE: 10/18/07

# ***ENTERED***

## SEQUENCE LISTING

### (1) GENERAL INFORMATION:

- (i) APPLICANT: Choulika, Andre  
Perrin, Arnaud  
Dujon, Bernard  
Nicolas, Jean-Francois
- (ii) TITLE OF INVENTION: Nucleotide Sequence Encoding the Enzyme  
I-SCEI and the Uses Thereof
- (iii) NUMBER OF SEQUENCES: 52
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
Dunner
  - (B) STREET: 1300 I Street, N.W.
  - (C) CITY: Washington
  - (D) STATE: D.C.
  - (E) COUNTRY: USA
  - (F) ZIP: 20005-3315
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: 10820843
  - (B) FILING DATE: 2004-04-09
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: 08336241
  - (B) FILING DATE: 07-NOV-1994
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 07/971,160
  - (B) FILING DATE: 05-NOV-1992
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 07/879,689
  - (B) FILING DATE: 05-MAY-1992
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Potter, Jane E.R.
  - (B) REGISTRATION NUMBER: 33,332
  - (C) REFERENCE/DOCKET NUMBER: 03495-0111-03000
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 202-408-4000
  - (B) TELEFAX: 202-408-4400

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 714 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGCATATGA AAAACATCAA AAAAAACCAG GTAATGAACC TCGGTCCGAA CTCTAAACTG  
60  
CTGAAAGAAT ACAAATCCCA GCTGATCGAA CTGAACATCG AACAGTTCGA AGCAGGTATC  
120  
GGTCTGATCC TGGGTGATGC TTACATCCGT TCTCGTGATG AAGGTAAAC CTACTGTATG  
180  
CAGTTCGAGT GGAAAAACAA AGCATACATG GACCACGTAT GTCTGCTGTA CGATCAGTGG  
240  
GTACTGTCCC CGCCGCACAA AAAAGAACGT GTTAACCACC TGGGTAACCT GGTAATCACC  
300  
TGGGGCGCCC AGACTTTCAA ACACCAAGCT TTCAACAAAC TGGCTAACCT GTTCATCGTT  
360  
AACAACAAAA AAACCATCCC GAACAACCTG GTTGAAACT ACCTGACCCC GATGTCTCTG  
420  
GCATACTGGT TCATGGATGA TGGTGGTAAA TGGGATTACA ACAAAAATC TACCAACAAA  
480  
TCGATCGTAC TGAACACCCA GTCTTTCACT TTCGAAGAAG TAGAATACCT GGTAAAGGGT  
540  
CTGCGTAACA AATTCCAAC TGAAGTTAC GTAAAAATCA AAAAAACAA ACCGATCATC  
600  
TACATCGATT CTATGTCTTA CCTGATCTTC TACAACCTGA TCAAACCGTA CCTGATCCCG  
660  
CAGATGATGT ACAAACAGCC GAACACTATC TCCTCCGAAA CTTTCCTGAA ATAA  
714

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 237 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	His	Met	Lys	Asn	Ile	Lys	Lys	Asn	Gln	Val	Met	Asn	Leu	Gly	Pro	
1				5					10					15		
Asn	Ser	Lys	Leu	Leu	Lys	Glu	Tyr	Lys	Ser	Gln	Leu	Ile	Glu	Leu	Asn	
			20					25					30			
Ile	Glu	Gln	Phe	Glu	Ala	Gly	Ile	Gly	Leu	Ile	Leu	Gly	Asp	Ala	Tyr	
		35					40					45				
Ile	Arg	Ser	Arg	Asp	Glu	Gly	Lys	Thr	Tyr	Cys	Met	Gln	Phe	Glu	Trp	
	50					55					60					
Lys	Asn	Lys	Ala	Tyr	Met	Asp	His	Val	Cys	Leu	Leu	Tyr	Asp	Gln	Trp	
65					70					75					80	
Val	Leu	Ser	Pro	Pro	His	Lys	Lys	Glu	Arg	Val	Asn	His	Leu	Gly	Asn	
				85					90					95		
Leu	Val	Ile	Thr	Trp	Gly	Ala	Gln	Thr	Phe	Lys	His	Gln	Ala	Phe	Asn	
			100					105					110			
Lys	Leu	Ala	Asn	Leu	Phe	Ile	Val	Asn	Asn	Lys	Lys	Thr	Ile	Pro	Asn	
		115					120					125				
Asn	Leu	Val	Glu	Asn	Tyr	Leu	Thr	Pro	Met	Ser	Leu	Ala	Tyr	Trp	Phe	
	130					135					140					
Met	Asp	Asp	Gly	Gly	Lys	Trp	Asp	Tyr	Asn	Lys	Asn	Ser	Thr	Asn	Lys	
145					150					155					160	
Ser	Ile	Val	Leu	Asn	Thr	Gln	Ser	Phe	Thr	Phe	Glu	Glu	Val	Glu	Tyr	
				165					170					175		
Leu	Val	Lys	Gly	Leu	Arg	Asn	Lys	Phe	Gln	Leu	Asn	Cys	Tyr	Val	Lys	
			180					185					190			
Ile	Asn	Lys	Asn	Lys	Pro	Ile	Ile	Tyr	Ile	Asp	Ser	Met	Ser	Tyr	Leu	
		195				200						205				
Ile	Phe	Tyr	Asn	Leu	Ile	Lys	Pro	Tyr	Leu	Ile	Pro	Gln	Met	Met	Tyr	
	210					215					220					
Lys	Leu	Pro	Asn	Thr	Ile	Ser	Ser	Glu	Thr	Phe	Leu	Lys				
225					230					235						

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 722 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AAAAATAAAA TCATATGAAA AATATTAAAA AAAATCAAGT AATCAATCTC GGTCCTATTT  
60  
CTAAATTATT AAAAGAATAT AAATCACAAT TAATTGAATT AAATATTGAA CAATTTGAAG  
120  
CAGGTATTGG TTTAATTTTA GGAGATGCTT ATATTCGTAG TCGTGATGAA GGTAAAACTT  
180  
ATTGTATGCA ATTTGAGTGG AAAAATAAGG CATACATGGA TCATGTATGT TTATTATATG  
240  
ATCAATGGGT ATTATCACCT CCTCATAAAA AAGAAAGAGT TAATCATTTA GGTAATTTAG  
300  
TAATTACCTG GGGAGCTCAA ACTTTTAAAC ATCAAGCTTT TAATAAATTA GCTAACTTAT  
360  
TTATTGTAAA TAATAAAAAA CTTATTCCTA ATAATTTAGT TGAAAATTAT TTAACACCTA  
420  
TGAGTCTGGC ATATTGGTTT ATGGATGATG GAGGTAAATG GGATTATAAT AAAAATTCTC  
480  
TTAATAAAG TATTGTATTA AATACACAAA GTTTTACTTT TGAAGAAGTA GAATATTTAC  
540  
TTAAAGGTTT AAGAAATAAA TTTCAATTAA ATTGTTATGT TAAATTAAT AAAAATAAAC  
600  
CAATTATTTA TATTGATTCT ATGAGTTATC TGATTTTTTA TAATTTAATT AAACCTTATT  
660  
TAATTCCTCA AATGATGTAT AAACGCCTA ATACTATTTT ATCCGAAACT TTTTAAAAAT  
720  
AA  
722

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 235 amino acids  
    (B) TYPE: amino acid  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Lys	Asn	Ile	Lys	Lys	Asn	Gln	Val	Met	Asn	Leu	Gly	Pro	Asn	Ser
1				5					10					15	
Lys	Leu	Leu	Lys	Glu	Tyr	Lys	Ser	Gln	Leu	Ile	Glu	Leu	Asn	Ile	Glu
			20					25					30		
Gln	Phe	Glu	Ala	Gly	Ile	Gly	Leu	Ile	Leu	Gly	Asp	Ala	Tyr	Ile	Arg

35					40					45					
Ser	Arg	Asp	Glu	Gly	Lys	Thr	Tyr	Cys	Met	Gln	Phe	Glu	Trp	Lys	Asn
50						55					60				
Lys	Ala	Tyr	Met	Asp	His	Val	Cys	Leu	Leu	Tyr	Asp	Gln	Trp	Val	Leu
65					70					75					80
Ser	Pro	Pro	His	Lys	Lys	Glu	Arg	Val	Asn	His	Leu	Gly	Asn	Leu	Val
				85					90					95	
Ile	Thr	Trp	Gly	Ala	Gln	Thr	Phe	Lys	His	Gln	Ala	Phe	Asn	Lys	Leu
			100					105					110		
Ala	Asn	Leu	Phe	Ile	Val	Asn	Asn	Lys	Lys	Leu	Ile	Pro	Asn	Asn	Leu
			115				120					125			
Val	Glu	Asn	Tyr	Leu	Thr	Pro	Met	Ser	Leu	Ala	Tyr	Trp	Phe	Met	Asp
	130					135					140				
Asp	Gly	Gly	Lys	Trp	Asp	Tyr	Asn	Lys	Asn	Ser	Leu	Asn	Lys	Ser	Ile
145						150					155				160
Val	Leu	Asn	Thr	Gln	Ser	Phe	Thr	Phe	Glu	Glu	Val	Cys	Tyr	Leu	Val
				165					170					175	
Lys	Gly	Leu	Arg	Asn	Lys	Phe	Gln	Leu	Asn	Cys	Tyr	Val	Lys	Ile	Asn
			180					185					190		
Lys	Asn	Lys	Pro	Ile	Ile	Tyr	Ile	Asp	Ser	Met	Ser	Tyr	Leu	Ile	Phe
	195					200					205				
Tyr	Asn	Ile	Ile	Lys	Pro	Tyr	Leu	Ile	Pro	Gln	Met	Met	Tyr	Lys	Leu
	210					215					220				
Pro	Asn	Thr	Ile	Ser	Ser	Glu	Thr	Phe	Leu	Lys					
225						230				235					

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 754 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCGGATCCAT GCATATGAAA AACATCAAAA AAAACCAGGT AATGAACCTG GGTCCGAAC T  
60

CTAAACTGCT GAAAGAATAC AAATCCCAGC TGATCGAACT GAACATCGAA CAGTTCGAAG  
120

CAGGTATCGG TCTGATCCTG GGTGATGCTT ACATCCGTTC TCGTGATGAA GGTAAACCT  
180